THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 290, NO. 5, pp. 2617–2629, January 30, 2015

Structural and Thermodynamic Insights into Chitooligosaccharide Binding to Human Cartilage Chitinase 3-like Protein 2 (CHI3L2 or YKL-39)*

Received for publication, June 11, 2014, and in revised form, November 20, 2014 Published, JBC Papers in Press, December 4, 2014, DOI 10.1074/jbc.M114.588905

Araya Ranok[‡], **Jantana Wongsantichon**[§1], **Robert C. Robinson**[§¶1,2], **and Wipa Suginta**[‡3]

From the [‡]*Biochemistry-Electrochemistry Research Unit, School of Biochemistry, Institute of Science, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand, the* [§]*Institute of Molecular and Cell Biology, Agency for Science, Technology, and Research (A*STAR), Biopolis Drive, Singapore 138673, Singapore, and the* [¶]*Department of Biochemistry, National University of Singapore, 8 Medical Drive, Singapore 117597, Singapore*

Background: Human YKL-39 is currently recognized as a biomarker for osteoarthritis.
Results: Crystal structures of YKL-39 reveal that chitooligosaccharide induces local conformational changes to stabilize sugar·protein complexes and that the protein contains five binding subsites for sugars.
Conclusion: YKL-39 binds to chitooligosaccharide through enthalpic reactions.
Significance: Our findings suggest how YKL-39 interacts with GlcNAc moieties of the natural ligands, which may possibly activate local tissue inflammation.

Four crystal structures of human YKL-39 were solved in the absence and presence of chitooligosaccharides. The structure of YKL-39 comprises a major $(\beta/\alpha)_8$ triose-phosphate isomerase barrel domain and a small $\alpha + \beta$ insertion domain. Structural analysis demonstrates that YKL-39 interacts with chitooligosaccharides through hydrogen bonds and hydrophobic interactions. The binding of chitin fragments induces local conformational changes that facilitate tight binding. Compared with other GH-18 members, YKL-39 has the least extended chitin-binding cleft, containing five subsites for sugars, namely (−3)(−2)(−1)(+1)(+2), with Trp-360 playing a prominent role in the sugar-protein interactions at the center of the chitin-binding cleft. Evaluation of binding affinities obtained from isothermal titration calorimetry and intrinsic fluorescence spectroscopy suggests that YKL-39 binds to chitooligosaccharides with K_d values in the micromolar concentration range and that the binding energies increase with the chain length. There were no significant differences between the K_d values of chitopentaose and chitohexaose, supporting the structural evidence for the five binding subsite topology. Thermodynamic analysis indicates that binding of chitooligosaccharide to YKL-39 is mainly driven by enthalpy.

Human chitinases and chitinase-like proteins (CLPs)[4] are members of the glycosyl hydrolase family 18 (GH-18), involved in several physiological processes such as tissue remodeling, injury, and inflammation (1–4). Although human chitotriosidase (CHIT1) and acidic mammalian chitinase are active enzymes that hydrolyze natural chitins (5–8), CLPs are nonenzymatic homologs that possess the chitin-binding groove on the surface of their $(\beta/\alpha)_8$ TIM barrel domains, which permits these proteins to bind to chitooligosaccharides with high affinity (9–11). CLPs, including chitinase 3-like 1 protein (CHI3L1 or YKL-40 or HC gp-39) (12), chitinase 3-like 2 protein (CHI3L2 or YKL-39) (13), oviduct-specific glycoprotein (oviductin or Mucin9) (14), and stabilin-1 interacting CLPs (15), lack chitinase activity because of the substitution of an essential catalytic residue (glutamic acid) at the end of the D*XX*D*X*D*X*E conserved motif with either leucine, isoleucine, or tryptophan (1).

Because YKL-39 is closely related in size and sequence to YKL-40, it was named following the convention for that homolog, which is based on the following three N-terminal amino acid residues, tyrosine (Tyr), lysine (Lys), and leucine (Leu), and an apparent molecular mass of 39 kDa. YKL-39 is secreted from articular chondrocytes (13). YKL-39 mRNA has been detected in lung, heart, and glioblastoma but not in brain, spleen, or pancreas (13, 16). YKL-39 mRNA was also detected in macrophages that were strongly stimulated by a combination of IL-4 and TGF-β (17). YKL-39 is currently recognized as a specific biomarker for the activation of chondrocytes and for the progress of osteoarthritis (18–20), a degenerative joint disease involving the degradation of articular cartilage and subchondral bone that globally affects 25% of adults aged over 65 years (21). Real time PCR and DNA microarray analyses showed that

* This work was supported by the Office of the Higher Education Commission through CHE333 PhD-THA-SUP Scholarship Grant CHE500307 and Suranaree University of Technology, Thailand, Grant SUT1-102-57-24-19.

The atomic coordinates and structure factors (codes 4P8U, 4P8V, 4P8W, and 4P8X) have been deposited in the Protein Data Bank (http://wwpdb.org/).

[1] Supported by the Biomedical Research Council of the Agency for Science, Technology, and Research (A*STAR).

[2] To whom correspondence may be addressed: Institute of Molecular and Cell Biology, Proteos, 61 Biopolis Dr., Singapore 138673, Singapore. Fax: 65-6586-9832; E-mail: rrobinson@imcb.a-star.edu.sg.

[3] To whom correspondence may be addressed: Suranaree University of Technology, 111 University Ave., Nakhon Ratchasima 30000, Thailand. Fax: 66-44-224185; E-mail: wipa@sut.ac.th.

[4] The abbreviations used are: CLP, chitinase-like protein; CHIT1, chitotriosidase; ITC, isothermal titration calorimetry; BisTris, 2-[bis(2-hydroxyethyl)-amino]-2-(hydroxymethyl)propane-1,3-diol; Trx, thioredoxin; PDB, Protein Data Bank; TIM, triose-phosphate isomerase.

YKL-39 mRNA was significantly up-regulated in the cartilage of patients with advanced osteoarthritis (19). Moreover, the level of YKL-39 mRNA expression was positively correlated with collagen type 2 up-regulation in both early and late stages of the disease (20). YKL-39 was also found to induce an autoimmune response in patients with rheumatoid arthritis (22, 23), as well as in a rheumatoid arthritis mouse model (24). A recent study in human embryonic kidney (HEK293) and human glioblastoma (U87 MG) cells showed that YKL-39-activated signal transduction was regulated through the phosphorylation of ERK1/ERK2 kinases (25). YKL-39 was later reported to enhance cell proliferation, colony formation, and type II collagen expression in mouse chondrogenic ATDC5 cells (26). It may therefore act as a novel growth/differentiation factor for articular cartilage chondrocytes, which regulate joint homeostasis in adults. However, the mechanistic details of how YKL-39 regulates cell proliferation and cell differentiation remain to be identified.

Previously, the crystal structure of an N35Q mutant of YKL-39, bound to $GlcNAc_6$, was reported (27). In the work reported here, we employ both crystallographic and thermodynamic studies to evaluate the binding of chitooligosaccharides of different lengths to YKL-39. Inspection of the crystal structures of YKL-39 in the absence and presence of chitooligosaccharides clearly indicates that YKL-39 binds specifically to chitooligosaccharides and that the binding strength is dependent on the length of the chitin chain. These structural data are discussed in relation to binding affinity, subsite topology, and the thermodynamic contributions to the sugar-protein interactions.

MATERIALS AND METHODS

Gene Cloning—The nucleotide sequence of the full-length *CHI3L2* gene encoding chitinase 3-like protein 2 or YKL-39 was retrieved from the GenBank™ database (accession number NM_004000), and the gene was amplified from a human cDNA template by the PCR technique (GeneScript Corp.) and cloned into the pET32a(+) expression vector. The recombinant YKL-39 was expressed as a fusion protein containing a cleavable thioredoxin (Trx) fragment, followed by a hexahistidine tag at the N terminus of the YKL-39 polypeptide (28). The Trx fragment was fused to the protein to increase its solubility, and the His_6 tag was included to aid purification. Nucleotide sequences of both sense and antisense strands of the *CHI3L2* fragment were confirmed by automated DNA sequencing (First Base Laboratories, Malaysia).

Recombinant Protein Expression and Purification—Recombinant YKL-39, lacking the 26-amino acid signal sequence, was expressed at high levels in *Escherichia coli* BL21 (DE3) (28). YKL-39-expressing cells were harvested by centrifugation, resuspended in lysis buffer (50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 1 mM PMSF, 1 $mg \cdot ml^{-1}$ lysozyme and 1% (v/v) Triton X-100), and then lysed on ice using a Sonopuls Ultrasonic homogenizer with a 20-mm diameter probe. The crude supernatant obtained after centrifugation at 19,000 rpm for 1 h was filtered through a 0.45-μm cutoff membrane filter (Millipore), and then applied to a HisTrap™ HP (1.0 × 5.0 cm) pre-packed column (GE Healthcare), connected to an ÄKTAprime™ Plus system (GE Healthcare). The column was equilibrated with 10 column volumes of equilibration buffer (50 mM Tris-HCl, pH 8.0, 50 mM NaCl), with a constant flow rate of 1 $ml \cdot min^{-1}$. After the column was thoroughly washed, bound protein was eluted with 10 column volumes of 250 mM imidazole, pH 8.0, and imidazole was then removed with a HiPrep™ 26/10 desalting column (GE Healthcare). The purified Trx/His_6/YKL-39 fusion protein was subsequently treated with enterokinase, following the manufacturer's instruction (GenScript), to remove the fusion tag. The Trx/His_6 segment was resolved from the YKL-39 polypeptide using a HisTrap™ HP (1.0 × 1.0 cm) prepacked column (GE Healthcare). Unbound fractions containing YKL-39 were pooled, concentrated, and then further purified to homogeneity using a HiLoad™ 16/60 Superdex™ 200 preparation grade gel filtration column (GE Healthcare). Fractions containing highly purified YKL-39 were combined and then exchanged into 20 mM Tris-HCl buffer, pH 8.0. The pooled fraction was then concentrated to 10–20 $mg \cdot ml^{-1}$ using a Vivaspin™ membrane concentrator. Protein concentrations were determined by the Pierce™ BCA assay (Novagen, Darmstadt, Germany). Aliquots of the purified YKL-39 were flash-frozen in liquid N_2 and then stored at −30 °C.

Site-directed Mutagenesis—The recombinant plasmid pET32a-(+)/CHI3L2 was used as DNA template in PCR-based site-directed mutagenesis. For the W36A mutant, the forward and reverse primers were, respectively, 3′-GTTTGCTACTTTACCAACGCATCCCAGGACCGGCAGGAACC-5′ and 5′-GGTTCCTGCCGGTCCTGGGATGCGTTGGTAAAGTAGCAAAC-3′. For the Y243A mutant, the forward and reverse primers were, respectively, 3′-GACAGAGGGCCAAGCTCCTACGCAAATGTGGATATGCTGTGGGG-5′ and 5′-CCCCACAGCATATTCCACATTTGCGTAGGAGCTTGGCCCTCTGTC-3′. For the W360A mutant, the forward and reverse primers were, respectively, 3′-CCTGGGAGGAGCCATGATCGCCTCTATTGACATGGATGAC-5′ and 5′-GTCATCCATGTCAATAGAGGCGATCATGGCTCCTCCCAGG-3′. The underlined sequences indicate the mutated codons. Site-directed mutagenesis was performed following the QuikChange site-directed mutagenesis protocol of Stratagene. The DpnI-treated DNA was transformed into *E. coli* XL1-Blue competent cells. The recombinant plasmids obtained from positive colonies were extracted using QuickClean II plasmid miniprep kits (GenScript, Piscataway, NJ) and were then re-transformed into *E. coli* DH5α cells. To verify that mutations were correct, the nucleotide sequences of the sense and antisense strands of the PCR fragment were determined by automated sequencing (First BASE Laboratories Sdn Bhn, Selangor Darul Ehsan, Malaysia). The mutant proteins were expressed and purified using the same protocol as for the wild-type protein.

Protein Crystallization—Initial crystallization screens were set up using a Screenmaker 96 + 8™ Xtal (Innovadyne Technologies Inc.) with sitting drop CrystalQuick™ plates (Greiner Bio-one, Germany). For each crystallization drop, 170 nl of freshly prepared YKL-39 (12.75 $mg \cdot ml^{-1}$ dissolved in 20 mM Tris-HCl, pH 8.0) was added to an equal volume of each precipitating agent from three screening kits, including Wizard I and II (Emerald BioSystems) and Crystal Screen HT™ (Hampton Research). Crystal optimization was performed using the hanging drop vapor diffusion method. After mixing equal vol-

umes of protein and mother liquor, hexagonal single crystals were observed after 1 day at 25 °C under the condition of 30% (w/v) PEG 3350, 0.2 M Li_2SO_4, 0.1 M BisTris, pH 5.5. The crystals were allowed to grow for 1 week. For crystal complexes, the YKL-39 crystals were soaked with chitooligosaccharides overnight at 25 °C, using optimized concentrations (0.1 mM for $GlcNAc_5$ and $GlcNAc_6$, 5 mM for $GlcNAc_4$, 10 mM for $GlcNAc_3$, and $GlcNAc_2$) prepared in the mother liquor and then flash-frozen in liquid nitrogen for subsequent x-ray diffraction analysis.

Data Collection, Processing, and Structure Determination—The YKL-39 crystals, either ligand-free or complexed with chitooligosaccharides, were exposed to 1.00 Å wavelength x-rays at the BL13B1 beamline, National Synchrotron Radiation Research Center, Taiwan. Data were collected on an ADSC Quantum 315 CCD detector. All diffraction data were indexed, integrated, and scaled using the program HKL2000 (29), and molecular replacement was employed to obtain phase information using the program MOLREP from the CCP4 suite (30). The structure of YKL-39 bound to $GlcNAc_2$ was solved using the previously published structure of YKL-39 in complex with $GlcNAc_6$ (PDB code 4AY1) as the search model (27). Other data sets, including ligand-free YKL-39 and YKL-39 in complex with $GlcNAc_4$ and $GlcNAc_6$, were solved using the final structure of the YKL-39·$GlcNAc_2$ complex as the model for rigid body refinement. The analyses of the electron density map $F_{obs} - F_{cal}$ and $2F_{obs} - F_{cal}$ and model building were carried out in COOT (31) and restrained refinement in REFMAC5 within CCP4 (32) and Phenix (33). The geometry of each final model was validated by PROCHECK (34). Evaluation of the secondary structure indicated no residues in the outlier regions of the Ramachandran plots. However, the amino acid residue at position 318 was found to be Trp instead of Arg. This discrepancy has been suggested to arise from a single nucleotide polymorphism, which is related to tissue specificity (35). The final $2F_{obs} - F_{cal}$ omit map, contoured at 1.0 σ, clearly showed the electron density maps for $GlcNAc_2$, $GlcNAc_4$, and $GlcNAc_6$ with full occupancy. The structures and electron density maps of all the refined structures were created and displayed by PyMOL (36) and LIGPLOT (37). Atomic coordinates and structure factors of the final models of YKL-39 have been deposited in the Protein Data Bank with PDB accession numbers 4P8U for apo-YKL-39, 4P8V for the YKL-39·$GlcNAc_2$ complex, 4P8W for the YKL-39·$GlcNAc_4$ complex, and 4P8X for the YKL-39·$GlcNAc_6$ complex.

Binding Study by Isothermal Titration Calorimetry (ITC)—Binding of chitooligosaccharides $GlcNAc_2$ to $GlcNAc_6$ to YKL-39 was investigated by ITC. This technique measures heat released or absorbed during the binding event, providing information about binding thermodynamics and yielding the stoichiometry (n), equilibrium binding association constant (K_a), enthalpy change (ΔH), Gibb's free energy (ΔG), and the entropy change (ΔS) of the reaction in a single experiment (38, 39). Experiments were performed at 25 °C. ITC experiments were carried out at least three times using the ITC-200 system (Microcal Inc) at 25 °C with a stirring speed of 260 rpm. For experiments with $GlcNAc_5$ and $GlcNAc_6$, 4 μl of 0.25 mM chitosugar was injected into the 300-μl calorimeter cell, containing 20 mM potassium phosphate buffer, pH 8.0, and 10 μM purified YKL-39. The injections were repeated 29 times over 140-s intervals. The background was measured by injecting the corresponding ligand into the cell containing only the buffer. Experiments with $GlcNAc_2$, $GlcNAc_3$, and $GlcNAc_4$ were performed as described above, but the concentration of each protein/sugar reaction was re-optimized as follows: 30 μM YKL-39 and 0.45 mM $GlcNAc_4$; 15 μM YKL-39 and 3 mM $GlcNAc_3$; and 15 μM YKL-39 and 4 mM $GlcNAc_2$. The ITC data were collected and analyzed using the Microcal Origin version 7.0 software. The ITC profile obtained by injecting the corresponding ligand into the reaction cell containing buffer without YKL-39 was subtracted from the corresponding data set. The resultant data were fitted by a single-site binding model in the nonlinear least square algorithm. The thermodynamic parameters, including binding stoichiometry (n), the equilibrium binding association constant (K_a), and the enthalpy change (ΔH) were subsequently evaluated. The Gibb's free energy (ΔG) and the entropy change (ΔS) were calculated from the relationship shown in Equation 1,

$$\Delta G = \Delta H - T\Delta S = -RT\ln(K_a) \quad \text{(Eq. 1)}$$

where R is the gas constant (1.98 cal·K^{-1} mol^{-1}) and T the absolute temperature in kelvin.

Binding Study by Fluorescence Spectroscopy—Each purified YKL-39 variant was titrated with different concentrations of the chitooligosaccharides in 20 mM Tris-HCl, pH 8.0, at 25 °C. Changes in intrinsic tryptophan fluorescence were monitored directly in an LS-50 fluorescence spectrometer (PerkinElmer Life Sciences). The excitation wavelength was set at 295 nm, and emission intensities were collected over 300–450 nm with excitation and emission slit widths of 5 nm. For the wild-type YKL-39 and the mutant Y243A, a fixed amount of protein (25 μg) was titrated with 100 mM $GlcNAc_2$, 10 mM $GlcNAc_{3,4}$, or 0.1 mM $GlcNAc_{5,6}$. Much higher concentrations of ligands were required for titrating the mutants W36A and W360A (100 mM $GlcNAc_{2,3,4}$ and 25 mM $GlcNAc_{5,6}$) were used. Each protein spectrum was corrected for the buffer spectrum. Binding curves were evaluated using a nonlinear regression function available in Prism version 5.0 (GraphPad Software), following the single-site binding model shown in Equation 2,

$$\Delta F = \frac{(F_{max} - F_{min}) \times (L_0)}{K_d + (L_0)} \quad \text{(Eq. 2)}$$

where ΔF is the difference between fluorescence intensity before and after titration with the sugar ligand; F_{max} refers to the maximum emission intensity; F_{min} is the minimum emission intensity; L_0 is the initial concentration of ligand; and K_d is the equilibrium dissociation constant (micromolar).

RESULTS AND DISCUSSION

Sequence Analysis—Human mature wild-type YKL-39, lacking the 26-amino acid signal sequence, was cloned and functionally expressed in *E. coli* as a Trx/His_6 fusion protein (28). The amino acid sequence of the recombinant YKL-39 was identical to the 390-amino acid sequence of CHI3L2 isoform 1 (identifier Q15782-4) reported in the UniProtKB/Swiss-Prot database, with the sole exception that the amino acid at position

TABLE 1
Statistics of data and structural refinement

	YKL-39, PDB 4P8U	YKL39/$(GlcNAc)_2$, PDB 4P8V	YKL39/$(GlcNAc)_4$, PDB 4P8W	YKL39/$(GlcNAc)_6$, PDB 4P8X
Data collection statistics				
Beamline	BL13B1	BL13B1	BL13B1	BL13B1
Wavelength (Å)	1.00	1.00	1.00	1.00
Space group	$P4_12_12$	$P4_12_12$	$P4_12_12$	$P4_12_12$
Unit cell parameters (Å)	$a = 71.6$	$a = 70.7$	$a = 71.4$	$a = 71.2$
	$b = 71.6$	$b = 70.7$	$b = 71.4$	$b = 71.2$
	$c = 140.7$	$c = 140.3$	c = 140.5	c = 141.0
Resolution range (Å)	26.3-2.40	19.6-1.64	19.8-1.87	24.8-2.48
Resolution outer shell (Å)	2.44-2.40	1.67-1.64	1.90-1.87	2.54-2.48
No. of unique reflections	13746	50934	26230	12630
No. of observed reflections	181110	528564	225642	90749
Completeness (%)	99.7 (99.9)	99.5 (99.4)	99.0 (87.8)	99.6 (94.6)
Average redundancy per shell	15.3 (15.8)	12.1 (10.4)	7.4 (5.8)	6.8 (5.3)
$I/\sigma(I)$	101.0 (6.4)	36.3 (3.5)	45.1 (2.8)	11.0 (3.1)
R_{merge} (%)	3.0 (46.9)	6.0 (48.4)	4.3 (45.5)	15.2 (34.7)
Refinement statistics				
R_{factor} (%)	21.6	17.3	18.6	17.6
R_{free} (%)	24.3	20.5	23.6	22.6
No. of residues in protein	365	365	364	365
No. of protein atoms	2896	2904	2919	2894
No. of ligand atoms		29	57	85
No. of water molecules	168	699	428	263
Mean *B*-factor				
Protein	66.6	14.9	23.7	48.6
Ligand		13.6	28.8	53.1
Water	79.5	33.0	34.6	49.9
Root mean square deviations				
Bond length	0.013	0.010	0.010	0.008
Bond angle	1.39	1.42	1.37	1.28
Ramachandran plot				
Residues in favored region (%)	96.1	98.6	97.8	97.8
Residues in allowed region (%)	3.9	1.4	2.2	2.2
Residues in outlier (%)	0.0	0.0	0.0	0.0

318 was Trp instead of Arg. The divergence of this amino acid has been suggested to arise from a genetic variation, which naturally occurs through a single nucleotide polymorphism that is tissue-specific (35). After enterokinase cleavage to remove the Trx/His_6 fragment, the recombinant YKL-39 contains seven extra N-terminal residues (AMADIGS), and the intact polypeptide has a predicted mass of 41.5 kDa.

Overall Structures of YKL-39 in the Absence and Presence of Ligands—Recombinant YKL-39, purified to homogeneity, was subjected to crystallization trials. This *E. coli* expressed human YKL-39 readily crystallized in PEG 3350. Initial crystallographic analysis showed that the YKL-39 crystals belong to the space group $P4_12_12$, with one molecule in the asymmetric unit. Four crystal structures of YKL-39 were determined, including the apo-form and the complexes with $GlcNAc_2$, $GlcNAc_4$, and $GlcNAc_6$. Table 1 summarizes the data collection and refinement statistics of the final models of the YKL-39 structures. YKL-39 in complex with $GlcNAc_2$ was refined against the highest resolution data at 1.53 Å. The overall structure of YKL-39 includes two conserved domains (Fig. 1*A*) that are found in all GH-18 chitinases and chitinase-like proteins. The major $(\beta/\alpha)_8$ TIM barrel domain (domain I) comprises eight parallel strands B1–B8 (*purple*), alternating with eight helices A1–A8 (*cyan*) as shown in Fig. 1*B*. It is noticeable that helices A1, A3, and A6 are broken and contain short helices: G1-1 in helix A1; G3-1 and G3-2 in helix A3; G6-1 and G6-2 in helix A6. The TIM barrel domains of GH-18 glycosyl hydrolases are known to interact specifically with chitin oligosaccharides (40–45). The second domain is termed the small $\alpha+\beta$ insertion domain (domain II) (Fig. 1*A*, *gray*) and is located between the tail of strand B7 and the start of helix A7. It is made up of six anti-parallel strands connected by a short helix that forms a typical greek key motif. The exact function of this domain is unknown, but certain amino acid residues in this domain help to stabilize the sugar·protein complex (41, 43, 46).

Detailed structural analysis revealed that YKL-39 contains two disulfide bonds in the TIM barrel domain (data not shown). The second is between Cys-308 of the TIM barrel domain and Cys-372 of the small insertion domain. Both bonds help maintain the structural integrity of the protein. In addition, six *cis* peptide bonds are observed, namely Glu-42–Pro-43, Ser-62–Phe-63, His-112–Pro-113, Ile-145–Tyr-146, Lys-220–Pro-221, and Trp-360–Ser-361, two of which (Ile-145–Tyr-146 and Trp-360–Ser-361) participate directly in ligand binding. The structures of YKL-39 in the presence of chitin fragments reveal the binding surface of the TIM barrel domain, which consists of a long deep groove with the approximate dimensions 35 Å (long) × 17 Å (wide) × 7 Å (deep). This groove has been reported to interact with chitin fragments in other GH-18 homologs (40–46). The shape of this groove is a crevice in which both ends are open, permitting chitooligosaccharides of different sizes to be accommodated and extended out from the groove in either direction. The narrowest part of the crevice points toward the center of the TIM barrel domain, a feature that favors a bent sugar chain.

Chitooligosaccharide Binding Induces Structural Movements—Superimposition of the apo-YKL-39 structure on the structures in complex with $GlcNAc_2$, $GlcNAc_4$, and $GlcNAc_6$ gives C_α root mean square deviations for 326 residues of 0.5, 0.4, and 0.4 Å, respectively. These values reflect some structural noniden-

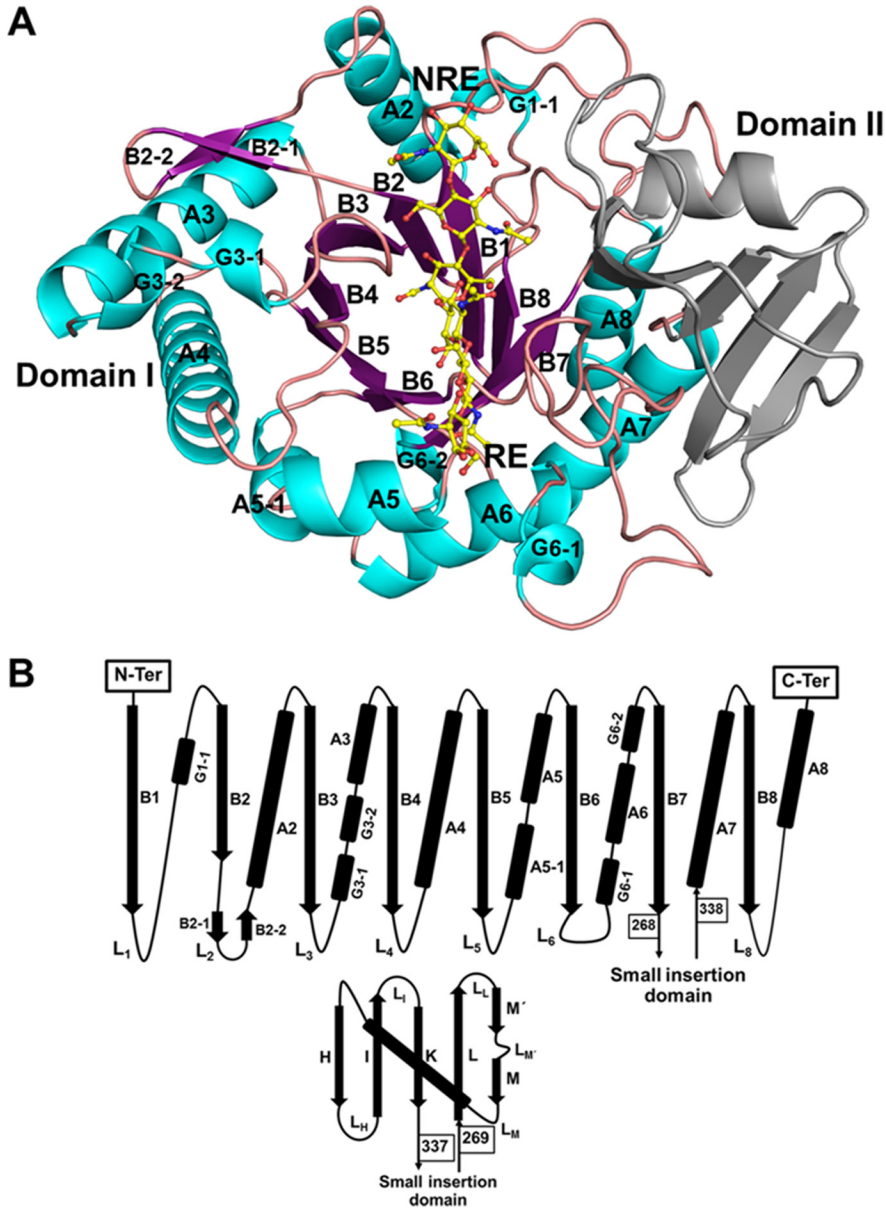

FIGURE 1. **Overall structure of YKL-39.** *A,* refined structure of YKL-39 in complex with $GlcNAc_6$. The chitin-binding cleft is located on the surface of the $(\beta/\alpha)_8$ TIM barrel domain (domain I) that is presented in *cyan* for helices and *purple* for strands, and the $\alpha+\beta$ insertion domain (domain II), flanked between strand B7 and helix A7 is shown in *gray*. The chitooligosaccharide $GlcNAc_6$ is shown as *ball* and *sticks* and colored by atoms with *yellow* for carbon, *blue* for nitrogen, and *red* for oxygen. *RE* and *NRE* denote the "reducing" and "nonreducing" ends of the sugar chain. *B,* structural topology of YKL-39. The eight strands are referred to as *B1* to *B8*, the loops as *L*, and the eight α-helices connecting the β-strands are referred to as *A1* to *A8*. The helices are often segmented. Short helices are referred to by the letter *G*, for instance *G1-1* indicating the short helix within the region of the A1 helix.

tity that may be induced upon sugar binding. Fig. 2*A* shows the $(\beta/\alpha)_8$ TIM barrel domains of unliganded and sugar-bound YKL-39, with small structural differences observed in the loop regions surrounding the surface of the chitin-binding cleft. Close inspection of the GlcNAc-binding subsites of the apoprotein (Fig. 2*A*, *blue*) in comparison with that of the protein in complex with the longest chitooligosaccharide $GlcNAc_6$ (Fig. 2, *cyan*) reveals considerable movements of loop L_1 on the surface of subsite −3, loops L_3 and $L_{M'}$ (see Fig. 1*B* for loop assignment) near subsites −2 and −1 and L_6 near subsites +2 and +3, in the structure with $GlcNAc_6$. As compared with the exterior of the native protein (Fig. 2*B*), four key residues were found to shift significantly toward the center of the chitin-binding cleft as follows: Tyr-104 and Leu-105 (part of the bottom loop L_3), and Met-364 and Phe-301 (part of the top loop $L_{M'}$). This narrows the cleft of the sugar-bound protein around subsites −2/−1 (Fig. 2*C*) by 1.3 Å. Such local movements engender close contacts between the GlcNAc rings and the binding residues around the corresponding subsites. In contrast, Tyr-243 (part of loop L_6) rotates away from its original position, widening the cleft at this particular subsite (+3) by 1.9 Å in comparison with that of the unliganded structure (Fig. 2*D*). This orientation of Tyr-243 increases access to subsite +2 from the subsite +3 direction, suggesting that longer sugars may be accommodated in this conformation. Taken together, these structural differences provide evidence that the binding of chitooligosaccharides induces local structural changes that strengthen the interactions between YKL-39 and chitosugars.

Specific Interactions of YKL-39 with Chitooligosaccharides—As in all GH-18 members, the sugar-binding cleft of YKL-39 comprises multiple subsites for GlcNAc units. The critical binding features of these subsites are aromatic residues, which bind the GlcNAc rings mainly through hydrophobic interactions, and polar amino acid side chains, which form hydrogen bonds with the saccharide chain. Inspection of the substrate-binding cleft of unliganded YKL-39 (Fig. 3*A*) and YKL-39 in complex with $GlcNAc_2$ (Fig. 3*B*), $GlcNAc_4$ (Fig. 3*C*), or $GlcNAc_6$ (Fig. 3*D*) reveals the binding behavior of individual chitooligosaccharides. $GlcNAc_2$ was found at the center of the binding cleft between subsites −2 and −1 (Fig. 3*B*). As summarized in Table 2, subsites −2 and −1 include a large number of binding residues. Residues Phe-63, Leu-105, Phe-301, Trp-360, and Met-364 interact with the GlcNAc ring at subsite −2, whereas Tyr-32, Ser-143, Tyr-104, Ile-145, Asp-213, Tyr-269, and Trp-360 form subsite −1. The residues around the center of the binding cleft make highly ordered interactions, which suggest that an arriving chitin chain would initially interact preferentially at the center of the binding groove. For a longer chitooligosaccharide, the interaction may become extended through occupation of the neighboring weaker affinity subsites, where the four sugar rings in $GlcNAc_4$ were extended from subsites −3 to +1 (Fig. 3*C*), and in the complex with $GlcNAc_6$ from subsites −3 to +3 (Fig. 3*D*). This affinity gradient is supported by the *B*-factor being lowest for −2GlcNAc (43.63 $Å^2$), indicating high rigidity through tight interactions at the internal site. Increases in *B*-factor for bound sugar at the extending subsites (−1GlcNAc, 44.76; +1GlcNAc, 44.75; +2GlcNAc, 52.48; +3GlcNAc, 81.64; and −3GlcNAc, 64.28 $Å^2$) suggest greater flexibility of the sugar rings, which make weak interactions with the binding residues at such subsites.

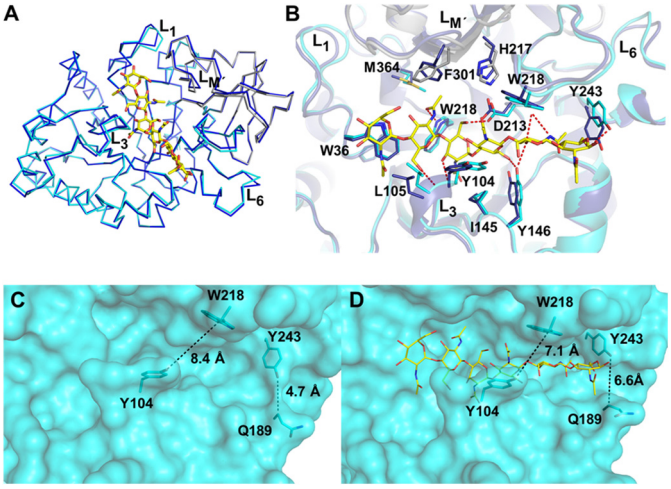

FIGURE 2. **Structural comparison of the YKL-39 structures in the absence and presence of $GlcNAc_6$.** *A,* structure of ligand-free YKL-39 (*royal blue*) was superimposed on the structure of YKL-39 in complex with $GlcNAc_6$ (*cyan* and *gray*). *B,* close up of the $GlcNAc_6$ binding pocket from *A*. The binding residues around the chitin-binding cleft that have different conformations in the two structures are included (Leu-105, Tyr-104, Met-364, Tyr-243, and Phe-301), and are displayed as *sticks*. *C,* surface representation of the sugar-binding groove of the unliganded form of YKL-39. *D,* surface representation of the sugar-binding groove of the YKL-39·$GlcNAc_6$ complex, showing local changes causing the closure of the surface area around subsites −2 and −1 and the widening of the area beyond subsite +2, relative to the binding groove in the apoprotein.

We further analyzed the hydrophobic interactions between chitohexaose $GlcNAc_6$ and aromatic/hydrophobic residues surrounding subsites −3 to +3 of the YKL-39-binding groove (Fig. 4*A*). Three residues, Trp-36 at subsite −3, Trp-360 at subsite −1, and Trp-218 at +2, were found to stack directly against the plane of the pyranose rings of the GlcNAc units, and these π-π interactions are expected to contribute significantly to binding at these subsites. The first two residues (Trp-36 and Trp-360) are completely conserved in other GH-18 homologs and act as key binding residues (42, 43, 45). Fig. 4*B* shows a number of hydrogen bonds that help to stabilize the sugar·protein complexes. Such hydrogen bonds are formed either directly between the sugar rings and the binding residues

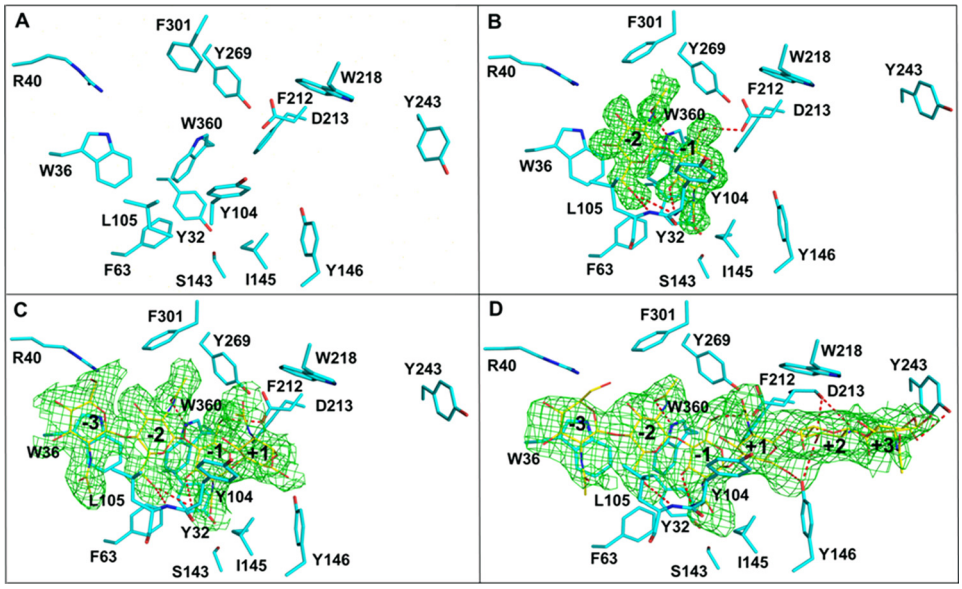

FIGURE 3. **Residues involved in chitooligosaccharide/YKL-39 binding.** The residues contributing hydrogen bonding and hydrophobic interactions are displayed as *sticks*. *A,* unliganded form of YKL-39; *B,* YKL-39 in complex with $GlcNAc_2$; *C,* YKL-39 in complex with $GlcNAc_4$; *D,* YKL39 in complex with $GlcNAc_6$. The sugar residues are shown as *sticks* and colored by atom, with *yellow* for carbon, *blue* for nitrogen, and *red* for oxygen. Hydrogen bonds are shown as *red dashed lines*. The $2F_o - F_c$ omit map is contoured at 1σ.

TABLE 2

Summary of the interactions of $(GlcNAc)_6$ with residues in each subsite of YKL-39, YKL-40 and hCHIT

Bold indicates stacking interaction residues. Bold with underlines indicates recognition residues. Bold and italic indicates hydrogen bonding residues.

Protein	Subsite −4	−3	−2	−1	+1	+2	+3
YKL-39		**Trp-36**	***Trp-360***	**Trp-360**		**Trp-218**	
			Leu-105	***Asp-213,Y104***	***Tyr-146***	***Tyr-146, Asp-213***	***Tyr-243***
		Arg-40	Phe-63, Phe-301, Met-364	Tyr-32, Ser-143, Ile-145, Tyr-269	Tyr-104, Leu-210, Phe-212, His-217	Phe-214	
YKL-40 (PDB codes 1HJW and 1NWT)		**Trp-31**	***Trp-352***	**Trp-352**	**Trp-99**		**Trp-212**
	Glu-70	***Asn-100, Glu-290***	***Asn-100, Glu-290***	***Asp-207, Arg-263***	***Tyr-141, Arg-263***	***Tyr-141***	
				Tyr-27, Ala-138, Leu-140, Tyr-206			
	Tyr-34	Arg-35, Trp-69	Tyr-27, Phe-58, Thr-269	Phe-261		Phe-208	
CHIT1(PDB code 1LG1)	Tyr-34	**Trp-31**	***Trp-358***	**Trp-358**	**Trp-99**	**Trp-218**	
			Asn-100, Glu-297	***Asp-138, Glu-140 Asp-213, Arg-269, Tyr-212***	***Tyr-141, Arg-269***		
		Arg-35	Tyr-27, Phe-58, Met-300	Tyr-267		Phe-214	

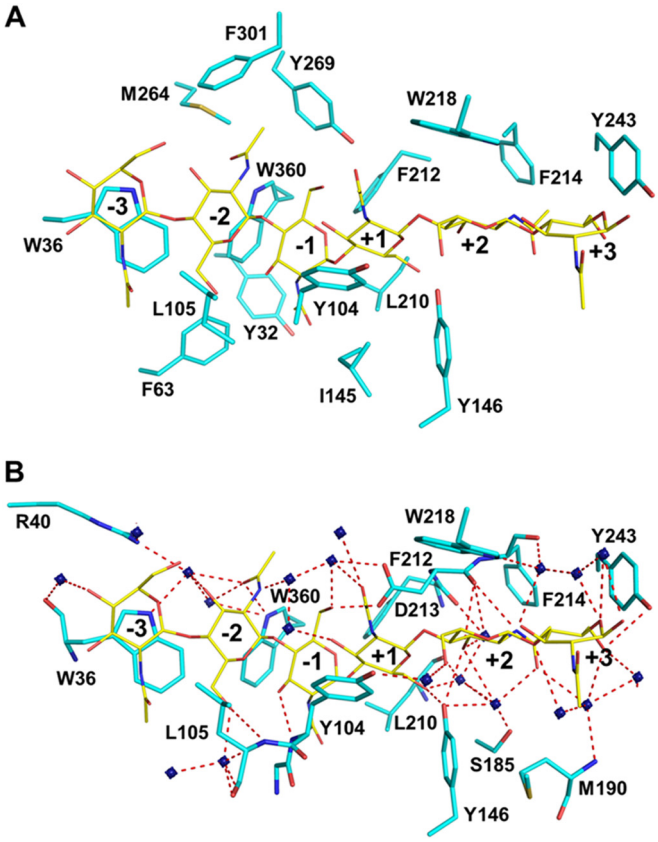

FIGURE 4. **Specific interactions within the chitin-binding cleft of YKL-39.** *A,* aromatic side chains that are involved hydrophobic interactions with the sugar molecule at each subsite. *B,* amino acid side chains and water molecules that form the hydrogen bonding network. Water molecules are shown as *blue diamonds* and hydrogen bonds are shown as *red dashed lines*.

or are mediated by water molecules. As summarized in Table 2, high densities of interactions are seen at subsites −2, −1, and +1, which are likely to determine the preferential binding at these sites.

Structural Comparison with Other Human GH-18 Members—Superimposition of the structure of the wild-type YKL39·$(GlcNAc)_6$ complex obtained in this study on that of YKL39·$GlcNAc_6$ complex (4AY1) reported by Schimpl *et al.* (27) results in a root mean square deviation of 0.22 for Cα positions over 334 atoms (Fig. 5). Thus, the two structures are essentially identical even though they are derived from crystals with different space groups. The differences between the two protein sequences are at residues 35 and 318.

In the structure by Schimpl *et al.* (27), Asn-35 was mutated to Gln to prevent glycosylation, generating mutant N35Q. In this study, YKL-39 was functionally expressed in *E. coli* as a native, nonglycosylated protein. In our structure, Asn-35 makes three salt bridges, two with Tyr-61 and Asp-75 and the third with a neighboring water molecule (Fig. 5*A*, *left panel, inset*). Gln-35 in Schimpl's structure flips to a vertical position and forms salt bridges with Tyr-61, Lys-74, and also a neighboring water molecule. Although residue 35 in both structures lies in subsite −3, it makes no contact with the GlcNAc ring at the corresponding subsite. Residue 318, Trp in native YKL-39, lies in a region of the protein distant from the chitin-binding groove. In our structure, Trp-318 forms hydrophobic interactions with Pro-325 and the stalk of Lys-340, although in structure 4AY1 Arg-318 forms a salt bridge with Asp-338 (Fig. 5*A*, *right panel, inset*). These are standard interactions between residues, in line with the view that these substitutions arise from naturally occurring polymorphisms. In our structure, all six GlcNAc rings could be fitted into subsites −3 to +3 (Fig. 5*B*), whereas only four GlcNAc rings of chitohexaose $GlcNAc_6$ were fitted into subsites −2 to +2 in the 4AY1 structure (Fig. 5*C*). The sugar rings found in both structures are well aligned in bent conformation. The −1GlcNAc rings of both sugar chains adopted the boat conformation, with the torsion angles of the C_α backbone being essentially identical.

The sugar-binding grooves of YKL-39, YKL-40 (42, 43), and CHIT1 (47) are very similar (Fig. 6*A*). Nevertheless, a few significant differences may explain divergent binding features between YKL-39 and other two homologs. The most obvious variation is at subsite −4. In CHIT1 and YKL-40, this subsite has Tyr-34 as the key residue forming a hydrophobic stacking interaction with −4GlcNAc. This residue is missing in YKL-39, and the closest residue, Asp-39, is unlikely to make productive contact with a GlcNAc ring (Table 2). As a result, no affinity at subsite −4 is likely in YKL-39. Another considerable difference is observed at subsite +3. In YKL-40, this subsite contains Trp-212, which stacks directly against the plane of +3GlcNAc. In

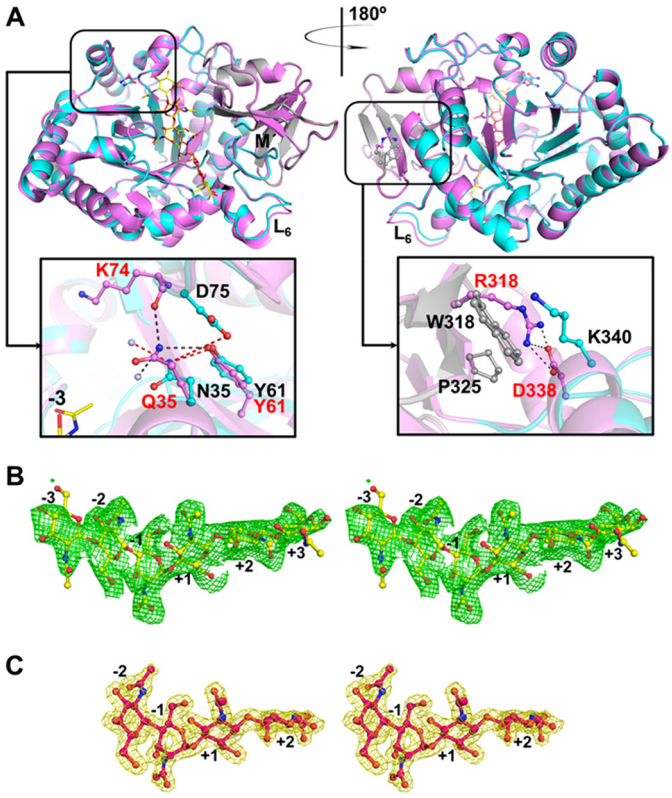

FIGURE 5. **Structural comparison of two YKL-39 structures.** *A*, superposition of the currently studied YKL-39 (PDB code 4P8X) and the previously reported YKL-39 (PDB code 4AY1, *magenta*) in complex with $GlcNAc_6$. The differences between the two protein sequences at residues N35Q (*left, inset*) and W318R (*right, inset*) are displayed as *ball* and *sticks*. *B*, stereo images of $GlcNAc_6$ within the binding cleft of the currently studied YKL-39 (*yellow*). *C*, stereo image of $GlcNAc_4$ of the previously reported YKL-39 (*red*). For both structures, nitrogen atoms are in *blue* and oxygen atoms in *red*. Both $2F_o - F_c$ omit maps are contoured at 1σ.

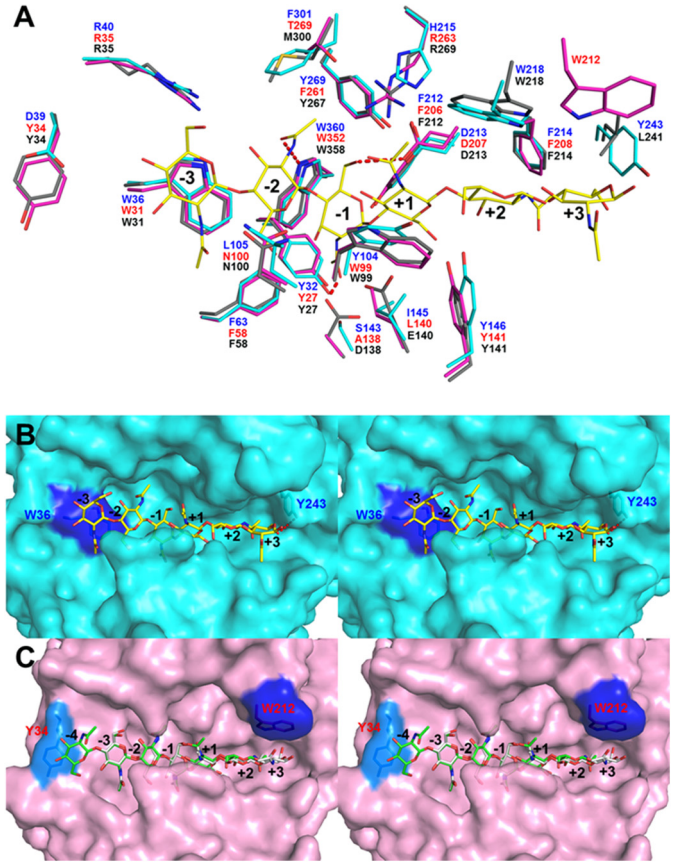

FIGURE 6. **Superimposition of YKL-39 and its human GH-18 homologs.** *A*, amino acid residues of YKL-39 (*cyan*), YKL-40 (*magenta*), and hCHIT (*gray*) that participate in sugar binding are shown. For the sake of clarity, only $GlcNAc_6$ in its complex with YKL-39 is shown (*yellow*). *B*, surface representation and of the sugar-binding cleft of YKL-39, showing chitohexaose occupying subsites −3 to +3. The interaction at subsite +3 appears weak. Trp-36 is highlighted in *dark blue* and Tyr-243 is labeled. *C*, surface representation of the sugar-binding cleft of YKL-40, showing six GlcNAc units of $GlcNAc_8$ occupying subsites −3 to +3 (PDB code 1NWU, *green*) or subsites −4 to +2 (PDB code 1HJW, *green*). Tyr-34 is highlighted in cyan and Trp-212 in *dark blue*.

contrast, Tyr-243 in YKL-39 forms a hydrogen bond at a distance of 4.5 Å with the equatorial C6-OH group of +3GlcNAc, suggesting that the interaction at this position is weaker than that in YKL-40. The −3GlcNAc unit at the nonreducing end of the $GlcNAc_6$ in YKL-39 (Fig. 6*B*) is well defined and interacts with Trp-36, whereas the +3GlcNAc unit is more flexible, adopting a more open position than the GlcNAc ring that forms a π-π stack with Trp-212 at subsite +3 in the YKL-40 structure (Fig. 6*C*). Weak affinity at the reducing end of the sugar chain is indicated by the large *B*-factor for +3GlcNAc (81.64 $Å^2$), reflecting high flexibility due to loose fitting at this location. Within the YKL-40 structure, the chitosugar occupies two positions between subsites −4 to +2 (Fig. 6*C*, *green*) and subsites −3 to +3 (Fig. 6*B*, *yellow*) due to the more extended binding cleft. The residue at the position homologous to that of the catalytic residue Glu-140 in the human chitinase (CHIT1) lies at bottom of subsite −1 and is Ile-145 in YKL-39 and Leu-140 in YKL-40 (Fig. 6*A*). Thus, neither YKL-39 nor YKL-40 have catalytic activity, although the bent conformation of the interacting sugar, rendering it susceptible to cleavage in active chitinases, is maintained in all the CLPs (41–43, 45, 47).

Binding Affinities and Five Major Subsite Topology—The enthalpic changes resulting from increasing chitooligosaccharide concentrations were measured using ITC, allowing the determination of binding affinities and inherent thermodynamic parameters for YKL-39 (39, 40). The heat release profiles of YKL-39 were measured during titration with discrete concentrations of each chitooligosaccharide (Fig. 7*A*, *upper panels*). Secondary plots of the injection peaks were integrated, yielding the enthalpy change as kilocalories/mol of injectant, and plotted as a function of the molar ratio of $GlcNAc_n$/YKL-39 (Fig. 7*A*, *lower panels*). All data obtained from the binding reactions were fitted using a single-site binding model with calculated stoichiometry (n) of 1.0, indicating that one molecule of the sugar interacts within the binding cleft of YKL-39. The suggested ratio of 1:1 of sugar/protein is consistent with that found in the crystal structures (Fig. 3, *B–D*), each of which shows only a single chitin chain. Theoretically, two or three molecules of $GlcNAc_2$ could bind within the chitin-binding groove of YKL-39, based on its size; however, only a single bound $GlcNAc_2$ is observed. This indicates that the subsites flanking the core subsites −2 and −1 have insufficient affinity to bind a second $GlcNAc_2$ molecule.

Analysis of the ITC data yields the equilibrium binding constants (K_d) for $GlcNAc_2$, $GlcNAc_3$, $GlcNAc_4$, $GlcNAc_5$, and $GlcNAc_6$, which are 204, 142, 1.7, 0.06, and 0.04 μM, respec-

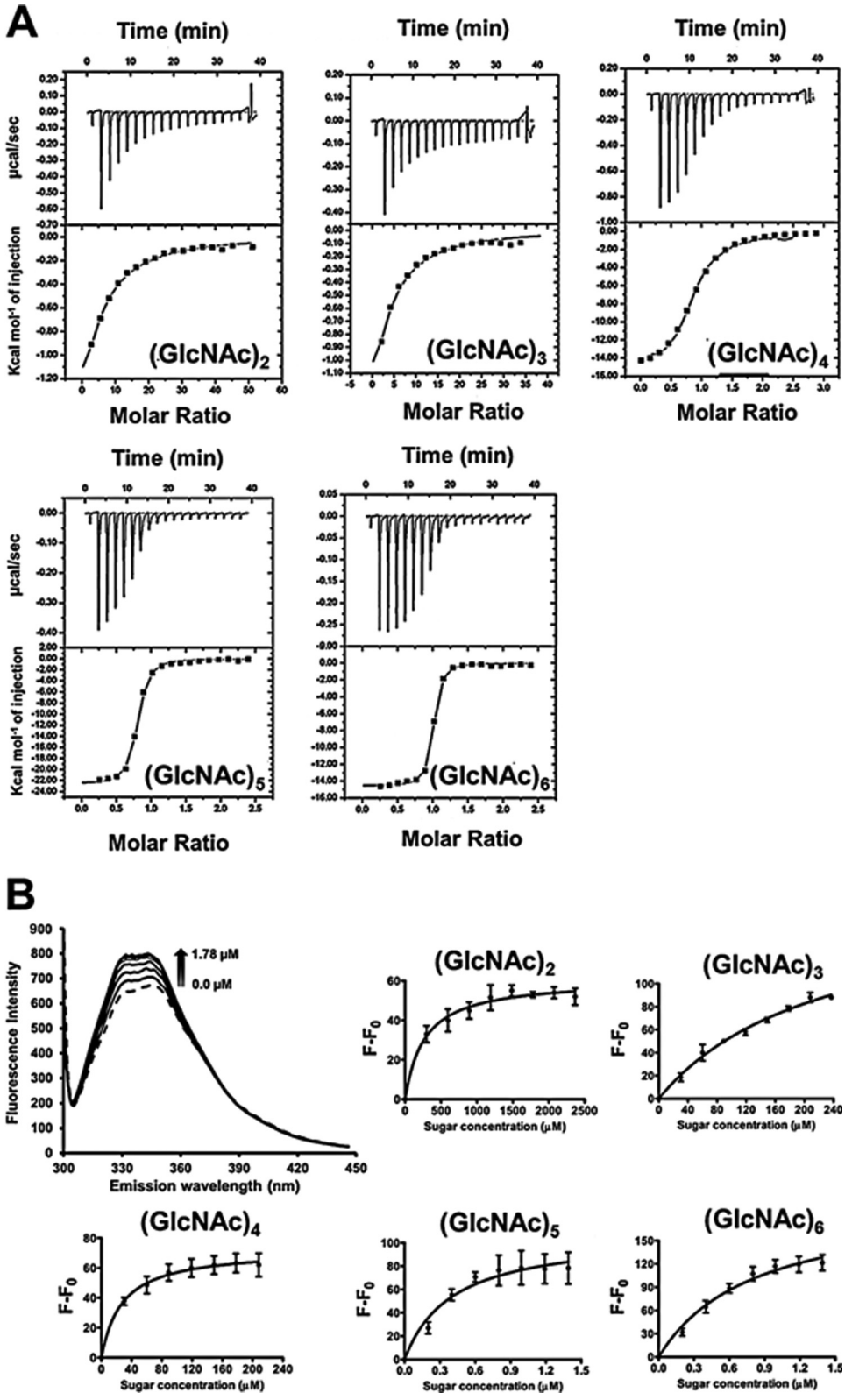

FIGURE 7. **YKL-39 binding to various chitooligosaccharides.** *A,* microcalorimetric titrations for the binding of chitooligosaccharide to YKL39 (*upper panels*) and binding isotherms with theoretical fittings (*lower panels*) obtained for the binding of $GlcNAc_2$ (*top left*), $GlcNAc_3$ (*top middle*), $GlcNAc_4$ (*top right*), $GlcNAc_5$ (*bottom left*), and $GlcNAc_6$ (*bottom right*) to YKL-39. Each ligand was added to protein solution in 20 mM potassium phosphate buffer, pH 8.0, at 25 °C. *B,* YKL-39 ligand binding studied by intrinsic fluorescence spectroscopy. YKL-39 was titrated with $GlcNAc_{2-6}$ as described in text. The fluorescence emission intensity increased with increasing concentrations of ligand, and emission spectra were collected at 300–450 nm (*top, left panel*). The excitation wavelength was set at 295 nm. The *dashed line* shows the emission spectrum of protein without ligand. Buffer emission spectra were subtracted in each case. The binding curves for $GlcNAc_2$ and $GlcNAc_3$, plotted as changes in relative fluorescence intensity against sugar concentration, are shown in the *top middle* and *right panels*, respectively. The binding curves for $GlcNAc_4$, $(GlcNAc)_5$, and $GlcNAc_6$ are presented in the *lower panels* (*left to right*) The binding curves were fitted with the nonlinear single site binding algorithm and the equilibrium dissociation constant (K_d) estimated from the concentration of ligand producing half-maximal fluorescence enhancement.

TABLE 3

Binding parameters of chitooligosaccharide binding to YKL-39 derived from isothermal microcalorimetry

The best fit parameters, including the stoichiometry of binding (n), the equilibrium association constant (K_a), and the enthalpy change (ΔH), were estimated from nonlinear regression, using a single-site binding model, of the binding thermograms shown in Fig. 7*A* The equilibrium dissociation constant (K_d) was an inverse value of K_a. The Gibb's free energy (ΔG) and the entropy change (ΔS) were calculated from Equation 1 as described in text. Values presented are means ± S.D. obtained from experiments carried out in triplicate. The ITC experiments were carried out at 25 °C (298 K).

YKL-39/chitooligosaccharide complex	Isothermal microcalorimetric (ITC) parameter					
	n	K_d	ΔH	ΔS	$T\Delta S$	ΔG
		μM	*kcal/mol*	*kcal/mol*	*kcal/mol*	*kcal/mol*
$(GlcNAc)_2$	1.0 ± 0.0	204 ± 3	−13.1 ± 4.4	−27.0 ± 14.8	7.6 ± 3.8	−5.1 ± 0.0
$(GlcNAc)_3$	1.0 ± 0.0	142 ± 8	−11.3 ± 1.3	−20.4 ± 4.5	6.1 ± 1.3	−5.2 ± 0.0
$(GlcNAc)_4$	0.9 ± 0.1	1.7 ± 0.7	−17.0 ± 2.0	−30.5 ± 5.9	9.1 ± 1.8	−7.9 ± 0.2
$(GlcNAc)_5$	0.9 ± 0.1	0.06 ± 0.04	−17.2 ± 3.7	−24.3 ± 13.3	7.2 ± 3.9	−10.0 ± 0.4
$(GlcNAc)_6$	1.1 ± 0.2	0.04 ± 0.02	−12.2 ± 1.5	−6.9 ± 5.2	2.0 ± 1.5	−10.2 ± 0.3

tively (Table 3). The decrease in the K_d value clearly indicates increasing binding affinity with increasing chitooligosaccharide length. These binding parameters are consistent with the structural data, which show that $GlcNAc_2$ only partially occupies the binding cleft at subsites −2 and −1 (Fig. 3*B*), whereas $GlcNAc_4$ occupies subsites −3 to +1. $GlcNAc_6$ stretches along the entire binding groove, occupying the six identified subsites and having the most stable binding of the chitooligosaccharides studied. Although the K_d values decrease with increasing chain length for $GlcNAc_2$, $GlcNAc_3$, $GlcNAc_4$, and $GlcNAc_5$, those of $GlcNAc_5$ (0.06 μM) and $GlcNAc_6$ (0.04 μM) are about equal, within the standard deviations (Table 3, ITC). The structure of the complex of YKL-39 with $GlcNAc_6$ (Fig. 4, *A* and *B*) suggests strong interactions from subsites −3 to +2, with Trp-36 forming a hydrophobic stack against −3GlcNAc and Trp-218 against +2GlcNAc, although binding at subsite +3 is weaker, because such interactions are absent. These binding characteristics were verified by data obtained from fluorescence spectroscopy. Changes in the fluorescence intensities of YKL-39/oligosaccharide solutions were measured at the maximum emission wavelength of 340 nm, with excitation at 295 nm. The fluorescence intensities were found to be dependent on the oligosaccharide concentration, indicating specific binding between the oligosaccharides and the protein. One example of this progressive fluorescence enhancement on binding of $GlcNAc_6$ is shown as a representative fluorescence profile (Fig. 7*B*, *left top panel*), which was analyzed and transformed to binding isotherms based on Equation 2. The fluorescence intensity data for each chitooligosaccharide were fitted reasonably well by the single-site binding model of a nonlinear regression function (Fig. 7*B*, *middle* and *right top panels* for $GlcNAc_2$ and $GlcNAc_3$, respectively, and *lower panel* from *left to right* for $GlcNAc_4$, $GlcNAc_5$, and $GlcNAc_6$, respectively). The estimated K_d values of YKL-39 WT are summarized in Table 4. In good agreement with ITC data, the binding strengths are in the order $GlcNAc_6 \cong GlcNAc_5 > GlcNAc_4 > GlcNAc_3 > GlcNAc_2$.

The values of K_d obtained from fluorescence measurements are in general higher than those from ITC measurements. This may reflect different reporting of binding by the two methods. In the fluorescence assay particular tryptophan residues lining the chitin-binding cleft are quenched during titrations with sugar. However, all interactions are accounted for in the ITC assay. The difference in K_d values obtained by the two methods is small for short-chain chitooligosaccharides and becomes greater with increasing chain length and tighter binding. More importantly, the K_d values for $GlcNAc_2$ derived from ITC and fluorescence methods are slightly different. This indeed indicates that Trp-360 that stacks against the facet of two GlcNAc rings at subsites −2/−1 plays an exclusive role in the sugar-protein interactions at such central subsites.

We further determined the role of three aromatic residues, Trp-36, Tyr-243, and Trp-360, in binding ligand. Trp-36 is located at subsite −3, Trp-360 at subsites −2/−1, and Tyr-243 at subsite +3 of the protein-binding cleft (see Table 2 for the summary of interactions). These residues were mutated to alanine, generating single mutants W36A, Y243A, and W360A, respectively. The effects of mutation on binding affinity were accessed in a fluorescence quenching assay. The summary of their K_d values compared with the wild-type values are shown in Table 3. Mutation of Tyr-243 caused only slight increases in K_d for the long-chain sugars: $GlcNAc_4$, $GlcNAc_5$, and $GlcNAc_6$, and caused no change in the K_d values for $GlcNAc_2$ and $GlcNAc_3$. The results suggest that Tyr-243 at subsite +3 plays a minor role in protein-sugar interactions and had no great influence on the overall binding properties of YKL-39.

In marked contrast, mutations of Trp-36, and Trp-360 to Ala caused dramatic loss binding f the protein toward $GlcNAc_2$ and $GlcNAc_3$, because no binding was detected at 100 mM sugar. The titration was not performed at higher concentrations due to solubility problems. The K_d values of both mutants for $GlcNAc_4$, $GlcNAc_5$, and $GlcNAc_6$ considerably increased. Notably, W360A showed greater increases in the K_d values than W36A. When compared with the WT values, the K_d values for the mutant W360A were 64- and 102-fold increased for $GlcNAc_5$ and $GlcNAc_6$, respectively. These result suggested that Trp-360 is crucial for maintaining tight binding at the center of the sugar-binding cleft (Fig. 3, *B* and *C*). No binding was observed with $GlcNAc_2$ and $GlcNAc_3$ indicating that Trp-36 is also important for sugar-ligand interactions at subsite −3. Mutation of this residue perhaps affected binding of the sugar at the neighboring subsite −2. Nevertheless, mutation of Trp-36 exhibited less severe impact on $GlcNAc_4$, $GlcNAc_5$, and $GlcNAc_6$ binding, indicating that the hydrogen bond and hydrophobic networks formed by the remaining subsites are sufficiently robust to

TABLE 4

Binding parameters of chitooligosaccharide binding to YKL-39 mutants as derived from fluorescence quenching assay

The purified YKL-39 (25 μg) was titrated with a different range of concentrations, and changes in fluorescence emission intensity in the course of protein·ligand formation were monitored by fluorescence spectroscopy. The values are obtained from fluorescence quenching assay. The K_d values (in micromolar) are presented as means ± S.D. and were calculated from experiments carried out in triplicate.

		W36A		Y243A		W360A	
Chitooligosaccharide	Wild type, K_d	K_d	Fold increased	K_d	Fold increased	K_d	Fold increased
$(GlcNAc)_2$	227 ± 45	ND[a]		213 ± 98	0.9	ND	
$(GlcNAc)_3$	180 ± 2 (32 ± 4)[b]	ND		172 ± 4	1.0	ND	
$(GlcNAc)_4$	27 ± 5	125 ± 22	4.6	11 ± 3	0.4	92 ± 7	3.4
$(GlcNAc)_5$	0.8 ± 0.4	21 ± 7	26	0.9 ± 0.05	1.1	51 ± 4	64
$(GlcNAc)_6$	0.6 ± 0.1 (0.6 ± 0.1)[b]	23 ± 3	38	0.8 ± 0.09	1.4	61 ± 5	102

[a] ND represents no detectable change in fluorescence intensity up to the highest ligand concentration used in the titration. Note the highest concentration of each ligand was limited by its solubility in the selected buffered solution.

[b] Values presented in parentheses are those reported by Schimpl *et al.* (27).

hold the long-chain sugars within the binding cleft, albeit with lower affinity.

The enthalpy change (ΔH), entropy change ($-T\Delta S$), and free energy change (ΔG) obtained from the ITC experiments (Table 3) clearly demonstrate that all chitooligosaccharides bind to YKL-39 in spontaneous, exothermic reactions. This is indicated by negative ΔG values in a range of -5.1 to -10.2 kcal·mol^{-1} for the overall binding reactions. Similar values estimated for $GlcNAc_5$ and $GlcNAc_6$ suggest insignificant differences in the binding strengths for these two chitooligosaccharides (Fig. 7*A*). The results of these kinetic analyses support the crystal structure observations and suggest that YKL-39 has a less extended sugar-binding cleft than its closely related homologs, which is dominated by five recognition sites, namely $(-3)(-2)(-1)(+1)(+2)$, instead of the six subsites observed in CHIT1 and YKL-40 (42, 43, 47).

Analysis of Thermodynamic Parameters Involving Ligand Binding—A plot of the free energy change (ΔG) as the sum of negative ΔH and positive $-T\Delta S$ demonstrates how the enthalpic and entropic parameters contribute to the binding by YKL-39 of chitooligosaccharides of increasing chain lengths (Fig. 8*A*). The negative ΔG values obtained for all chitooligosaccharides are clearly influenced by the dominant negative ΔH values, ranging from -11.3 to -17.2 kcal·mol^{-1}, although the entropic term ($-T\Delta S$) is consistently unfavorable, with positive values ranging from $+2.0$ to $+9.1$ kcal·mol^{-1} (Fig. 8*B* and Table 3). This indicates that the chitooligosaccharide/YKL-39-binding reactions are mainly driven by enthalpic, rather than the entropic, factors. The dominant ΔH term represents the specificity and strength of interaction derived from hydrogen bonding and electrostatic interactions between the two components (48–52). A mechanistic basis for this phenomenon is provided by the crystal structures of the ligand·protein complexes, in which the sugar rings are held by varying numbers of hydrogen bonds (Fig. 4*B*). The largest enthalpy change measured was for $GlcNAc_5$, suggesting that sugar specificity is mediated by five binding subsites, which is consistent with the structural evidence.

Conclusions—This study provides structural and thermodynamic insights into the binding of chitooligosaccharides to human YKL-39, a specific biomarker for osteoarthritis. Four crystal structures of YKL-39, in the absence and presence of chitooligosaccharides $GlcNAc_{2-6}$, were solved. The

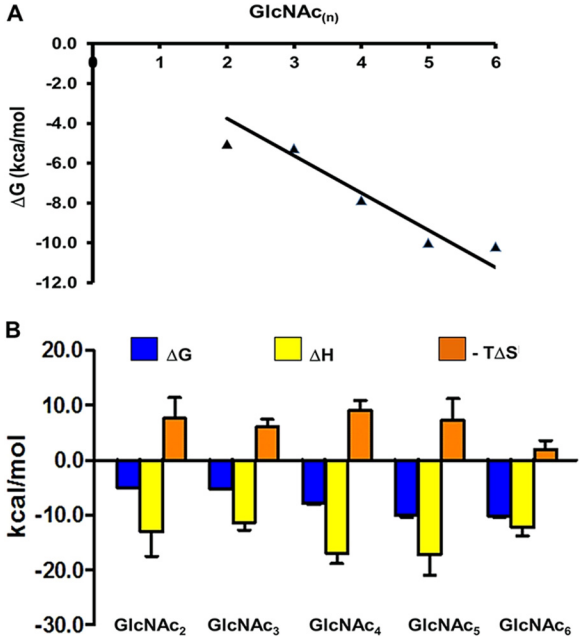

FIGURE 8. **Thermodynamic parameters of binding.** *A,* changes in the binding free energy obtained from Table 4 are plotted against the lengths of chitooligosaccharides. *B,* comparative plot of thermodynamic parameters (free energy, binding enthalpy, and entropy) for the binding of five chitooligosaccharides.

YKL-39 structure contains a major $(\alpha/\beta/)_8$ TIM barrel domain with a small insertion domain, similar to other GH-18 chitinases and CLPs. Superimposition of the crystal structures of ligand-free and ligand-bound YKL-39 suggests that binding of a chitooligosaccharide chain induces local conformational changes around the sugar-binding cleft that strengthen the sugar-protein interactions. The crystal structures of complexes with chitin fragments show that YKL-39 interacts with its sugar counterpart mainly through hydrophobic interactions, as well as a hydrogen bond network. ITC and fluorescence quenching data suggest that the protein binds chitooligosaccharides with very high affinity and that the binding strength increases with increasing length of the ligands. Both structural and thermodynamic evidence suggest that most of the binding is achieved with a five-subsite topology. Analysis of thermodynamic parameters suggests that all chitooligosaccharides bind to YKL-39 through enthalpy-driven reactions.

*Acknowledgments—We acknowledge the Biochemistry Laboratory, Center for Scientific and Technological Equipment, Suranaree University of Technology, for providing research facilities. Parts of this study were carried out on beamline BL13B1 at the National Synchrotron Radiation Research Center, Taiwan. We thank Joma Kanikadu Joy from Experimental Therapeutics Centre, A*STAR, for help with the ITC measurements. We also thank Dr. David Apps, Centre for Integrative Physiology, School of Biomedical Sciences, The University of Edinburgh, Edinburgh, Scotland, United Kingdom, for a critical proofreading of this manuscript.*

REFERENCES

1. Lee, C. G., Da Silva, C. A., Dela Cruz, C. S., Ahangari, F., Ma, B., Kang, M. J., He, C. H., Takyar, S., and Elias, J. A. (2011) Role of chitin and chitinase/chitinase-like proteins in inflammation, tissue remodeling, and injury. *Annu. Rev. Physiol.* **73,** 479–501
2. Kawada, M., Hachiya, Y., Arihiro, A., and Mizoguchi, E. (2007) Role of mammalian chitinases in inflammatory conditions. *Keio J. Med.* **56,** 21–27
3. Elias, J. A., Homer, R. J., Hamid, Q., and Lee, C. G. (2005) Chitinases and chitinase-like proteins in T(H)2 inflammation and asthma. *J. Allergy Clin. Immunol.* **116,** 497–500
4. Stevens, A. L., Wishnok, J. S., Chai, D. H., Grodzinsky, A. J., and Tannenbaum, S. R. (2008) A sodium dodecyl sulfate–polyacrylamide gel electrophoresis–liquid chromatography tandem mass spectrometry analysis of bovine cartilage tissue response to mechanical compression injury and the inflammatory cytokines tumor necrosis factor α and interleukin-1β. *Arthritis Rheum.* **58,** 489–500
5. Hollak, C. E., van Weely, S., van Oers, M. H., and Aerts, J. M. (1994) Marked elevation of plasma chitotriosidase activity. A novel hallmark of Gaucher disease. *J. Clin. Invest.* **93,** 1288–1292
6. Boot, R. G., Renkema, G. H., Strijland, A., van Zonneveld, A. J., and Aerts, J. M. (1995) Cloning of a cDNA encoding chitotriosidase, a human chitinase produced by macrophages. *J. Biol. Chem.* **270,** 26252–26256
7. Boot, R. G., Renkema, G. H., Verhoek, M., Strijland, A., Bliek, J., de Meulemeester, T. M., Mannens, M. M., and Aerts, J. M. (1998) The human chitotriosidase gene. Nature of inherited enzyme deficiency. *J. Biol. Chem.* **273,** 25680–25685
8. Boot, R. G., Blommaart, E. F., Swart, E., Ghauharali-van der Vlugt, K., Bijl, N., Moe, C., Place, A., and Aerts, J. M. (2001) Identification of a novel acidic mammalian chitinase distinct from chitotriosidase. *J. Biol. Chem.* **276,** 6770–6778
9. Sekine, T., Masuko-Hongo, K., Matsui, T., Asahara, H., Takigawa, M., Nishioka, K., and Kato, T. (2001) Recognition of YKL-39, a human cartilage related protein, as a target antigen in patients with rheumatoid arthritis. *Ann. Rheum. Dis.* **60,** 49–54
10. Arias, E. B., Verhage, H. G., and Jaffe, R. C. (1994) Complementary deoxyribonucleic acid cloning and molecular characterization of an estrogen-dependent human oviductal glycoprotein. *Biol. Reprod.* **51,** 685–694
11. Politz, O., Gratchev, A., McCourt, P. A., Schledzewski, K., Guillot, P., Johansson, S., Svineng, G., Franke, P., Kannicht, C., Kzhyshkowska, J., Longati, P., Velten, F. W., Johansson, S., and Goerdt, S. (2002) Stabilin-1 and -2 constitute a novel family of fasciclin-like hyaluronan receptor homologues. *Biochem. J.* **362,** 155–164
12. Hakala, B. E., White, C., and Recklies, A. D. (1993) Human cartilage gp-39, a major secretory product of articular chondrocytes and synovial cells, is a mammalian member of a chitinase protein family. *J. Biol. Chem.* **268,** 25803–25810
13. Hu, B., Trinh, K., Figueira, W. F., and Price, P. A. (1996) Isolation and sequence of a novel human chondrocyte protein related to mammalian members of the chitinase protein family. *J. Biol. Chem.* **271,** 19415–19420
14. Lapensée, L., Paquette, Y., and Bleau, G. (1997) Allelic polymorphism and chromosomal localization of the human oviductin gene (MUC9). *Fertil. Steril.* **68,** 702–708
15. Kzhyshkowska, J., Mamidi, S., Gratchev, A., Kremmer, E., Schmuttermaier, C., Krusell, L., Haus, G., Utikal, J., Schledzewski, K., Scholtze, J., and Goerdt, S. (2006) Novel stabilin-1 interacting chitinase-like protein (SI-CLP) is upregulated in alternatively activated macrophages and secreted via lysosomal pathway. *Blood* **107,** 3221–3228
16. Kavsan, V., Dmitrenko, V., Boyko, O., Filonenko, V., Avdeev, S., Areshkov, P., Marusyk, A., Malisheva, T., Rozumenko, V., and Zozulya, Y. (2008) Overexpression of YKL-39 gene in glial brain tumors. *Scholarly Research Exchange* 10.3814/2008/814849
17. Gratchev, A., Schmuttermaier, C., Mamidi, S., Gooi, L., Goerdt, S., and Kzhyshkowska, J. (2008) Expression of osteoarthritis marker YKL-39 is stimulated by transforming growth factor β (TGF-β) and IL-4 in differentiating macrophages. *Biomark. Insights* **3,** 39–44
18. De Ceuninck, F., Marcheteau, E., Berger, S., Caliez, A., Dumont, V., Raes, M., Anract, P., Leclerc, G., Boutin, J. A., and Ferry, G. (2005) Assessment of some tools for the characterization of the human osteoarthritic cartilage proteome. *J. Biomol Tech.* **16,** 256–265
19. Steck, E., Breit, S., Breusch, S. J., Axt, M., and Richter, W. (2002) Enhanced expression of the human chitinase 3-like 2 gene (YKL-39) but not chitinase 3-like 1 gene (YKL-40) in osteoarthritic cartilage. *Biochem. Biophys. Res. Commun.* **299,** 109–115
20. Knorr, T., Obermayr, F., Bartnik, E., Zien, A., and Aigner, T. (2003) YKL-39 (chitinase 3-like protein 2), but not YKL-40 (chitinase 3-like protein 1), is up regulated in osteoarthritic chondrocytes. *Ann. Rheum. Dis.* **62,** 995–998
21. Breedveld, F. C. (2004) Osteoarthritis–the impact of a serious disease. *Rheumatology* **43,** (Suppl. 1) i4–i8
22. Du, H., Masuko-Hongo, K., Nakamura, H., Xiang, Y., Bao, C. D., Wang, X. D., Chen, S. L., Nishioka, K., and Kato, T. (2005) The prevalence of autoantibodies against cartilage intermediate layer protein, YKL-39, osteopontin, and cyclic citrullinated peptide in patients with early stage knee osteoarthritis: evidence of a variety of autoimmune processes. *Rheumatol. Int.* **26,** 35–41
23. Tsuruha, J., Masuko-Hongo, K., Kato, T., Sakata, M., Nakamura, H., Sekine, T., Takigawa, M., and Nishioka, K. (2002) Autoimmunity against YKL-39, a human cartilage derived protein, in patients with osteoarthritis. *J. Rheumatol.* **29,** 1459–1466
24. Sakata, M., Masuko-Hongo, K., Tsuruha, J., Sekine, T., Nakamura, H., Takigawa, M., Nishioka, K., and Kato, T. (2002) YKL-39, a human cartilage-related protein, induces arthritis in mice. *Clin. Exp. Rheumatol.* **20,** 343–350
25. Areshkov, P. A., and Kavsan, V. M. (2010) Chitinase 3-like protein 2 (CHI3L2, YKL-39) activates phosphorylation of extracellular signal-regulated kinases ERK1/ERK2 in human embryonic kidney (HEK293) and human glioblastoma (U87 MG) cells. *Tsitol. Genet.* **44,** 3–9
26. Miyatake, K., Tsuji, K., Yamaga, M., Yamada, J., Matsukura, Y., Abula, K., Sekiya, I., and Muneta, T. (2013) Human YKL39 (chitinase 3-like protein 2), an osteoarthritis-associated gene, enhances proliferation and type II collagen expression in ATDC5 cells. *Biochem. Biophys. Res. Commun.* **431,** 52–57
27. Schimpl, M., Rush, C. L., Betou, M., Eggleston, I. M., Recklies, A. D., and van Aalten, D. M. (2012) Human YKL-39 is a pseudo-chitinase with retained chitooligosaccharide-binding properties. *Biochem. J.* **446,** 149–157
28. Ranok, A., Khunkaewla, P., and Suginta, W. (2013) Human cartilage chitinase 3-like protein 2: cloning, expression, and production of polyclonal and monoclonal antibodies for osteoarthritis detection and identification of potential binding partners. *Monoclon. Antib. Immunodiagn. Immunother.* **32,** 317–325
29. Otwinowski, Z., and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326
30. Vagin, A., and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr.* **30,** 1022–1025
31. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60,** 2126–2132
32. Collaborative Computational Project No. 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* **50,** 760–763
33. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) PHENIX: building new software for automated

crystallographic structure determination. *Acta Crystallogr. D. Biol. Crystallogr.* **58,** 1948–1954
34. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl Crystallogr.* **26,** 283–291
35. Gerhard, D. S., Wagner, L., Feingold, E. A., Shenmen, C. M., Grouse, L. H., Schuler, G., Klein, S. L., Old, S., Rasooly, R., Good, P., Guyer, M., Peck, A. M., Derge, J. G., Lipman, D., Collins, F. S., Jang, W., Sherry, S., Feolo, M., Misquitta, L., Lee, E., Rotmistrovsky, K., Greenhut, S. F., Schaefer, C. F., Buetow, K., Bonner, T. I., Haussler, D., Kent, J., Kiekhaus, M., Furey, T., Brent, M., Prange, C., Schreiber, K., Shapiro, N., Bhat, N. K., Hopkins, R. F., Hsie, F., Driscoll, T., Soares, M. B., Casavant, T. L., Scheetz, T. E., Brownstein, M. J., Usdin, T. B., Toshiyuki, S., Carninci, P., Piao, Y., Dudekula, D. B., Ko, M. S., Kawakami, K., Suzuki, Y., Sugano, S., Gruber, C. E., Smith, M. R., Simmons, B., Moore, T., Waterman, R., Johnson, S. L., Ruan, Y., Wei, C. L., Mathavan, S., Gunaratne, P. H., Wu, J., Garcia, A. M., Hulyk, S. W., Fuh, E., Yuan, Y., Sneed, A., Kowis, C., Hodgson, A., Muzny, D. M., McPherson, J., Gibbs, R. A., Fahey, J., Helton, E., Ketteman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madari, A., Young, A. C., Wetherby, K. D., Granite, S. J., Kwong, P. N., Brinkley, C. P., Pearson, R. L., Bouffard, G. G., Blakesly, R. W., Green, E. D., Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmutz, J., Myers, R. M., Butterfield, Y. S., Griffith, M., Griffith, O. L., Krzywinski, M. I., Liao, N., Morin, R., Palmquist, D., Petrescu, A. S., Skalska, U., Smailus, D. E., Stott, J. M., Schnerch, A., Schein, J. E., Jones, S. J., Holt, R. A., Baross, A., Marra, M. A., Clifton, S., Makowski, K. A., Bosak, S., Malek, J., and MGC Project Team. (2004) The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). *Genome Res.* **14,** 2121–2127
36. Delano, W. L. (2002) *The PyMOL Molecular Graphics System,* DeLano Scientific LLC, San Carlos, CA
37. Laskowski, R. A., and Swindells, M. B. (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* **51,** 2778–2786
38. Perozzo, R., Folkers, G., and Scapozza, L. (2004) Thermodynamics of protein-ligand interactions: history, presence, and future aspects. *J. Recept. Signal. Transduct. Res.* **24,** 1–52
39. Doyle, M. L. (1997) Characterization of binding interactions by isothermal titration calorimetry. *Curr. Opin. Biotechnol.* **8,** 31–35
40. van Aalten, D. M., Komander, D., Synstad, B., Gåseidnes, S., Peter, M. G., and Eijsink, V. G. (2001) Structural insights into the catalytic mechanism of a family 18 exo-chitinase. *Proc. Natl. Acad. Sci. U.S.A.* **98,** 8979–8984
41. Fusetti, F., von Moeller, H., Houston, D., Rozeboom, H. J., Dijkstra, B. W., Boot, R. G., Aerts, J. M., and van Aalten, D. M. (2002) Structure of human chitotriosidase. Implications for specific inhibitor design and function of mammalian chitinase-like lectins. *J. Biol. Chem.* **277,** 25537–25544
42. Fusetti, F., Pijning, T., Kalk, K. H., Bos, E., and Dijkstra, B. W. (2003) Crystal structure and carbohydrate-binding properties of the human cartilage glycoprotein-39. *J. Biol. Chem.* **278,** 37753–37760
43. Meng, G., Zhao, Y., Bai, X., Liu, Y., Green, T. J., Luo, M., and Zheng, X. (2010) Structure of human stabilin-1 interacting chitinase-like protein (SI-CLP) reveals a saccharide-binding cleft with lower sugar-binding selectivity. *J. Biol. Chem.* **285,** 39898–39904
44. Songsiriritthigul, C., Pantoom, S., Aguda, A. H., Robinson, R. C., and Suginta, W. (2008) Crystal structures of *Vibrio harveyi* chitinase A complexed with chitooligosaccharides: implications for the catalytic mechanism. *J. Struct. Biol.* **162,** 491–499
45. Olland, A. M., Strand, J., Presman, E., Czerwinski, R., Joseph-McCarthy, D., Krykbaev, R., Schlingmann, G., Chopra, R., Lin, L., Fleming, M., Kriz, R., Stahl, M., Somers, W., Fitz, L., and Mosyak, L. (2009) Triad of polar residues implicated in pH specificity of acidic mammalian chitinase. *Protein Sci.* **18,** 569–578
46. Houston, D. R., Recklies, A. D., Krupa, J. C., and van Aalten, D. M. (2003) Structure and ligand-induced conformational change of the 39-kDa glycoprotein from human articular chondrocytes. *J. Biol. Chem.* **278,** 30206–30212
47. Brameld, K. A., and Goddard, W. A. (1998) Substrate distortion to a boat conformation at subsite −1 is critical in the mechanism of family 18 chitinases. *J. Am. Chem. Soc.* **120,** 3571–3580
48. Baban, J., Fjeld, S., Sakuda, S., Eijsink, V. G., and Sørlie, M. (2010) The roles of three *Serratia marcescens* chitinases in chitin conversion are reflected in different thermodynamic signatures of allosamidin binding. *J. Phys. Chem. B* **114,** 6144–6149
49. Bouvignies, G., Bernadó, P., Meier, S., Cho, K., Grzesiek, S., Brüschweiler, R., and Blackledge, M. (2005) Identification of slow correlated motions in proteins using residual dipolar and hydrogen-bond scalar couplings. *Proc. Natl. Acad. Sci. U.S.A.* **102,** 13885–13890
50. Flint, J., Nurizzo, D., Harding, S. E., Longman, E., Davies, G. J., Gilbert, H. J., and Bolam, D. N. (2004) Ligand-mediated dimerization of a carbohydrate-binding module reveals a novel mechanism for protein-carbohydrate recognition. *J. Mol. Biol.* **337,** 417–426
51. Norberg, A. L., Karlsen, V., Hoell, I. A., Bakke, I., Eijsink, V. G., and Sørlie, M. (2010) Determination of substrate binding energies in individual subsites of a family 18 chitinase. *FEBS Lett.* **584,** 4581–4585
52. Takano, K., Yamagata, Y., Funahashi, J., Hioki, Y., Kuramitsu, S., and Yutani, K. (1999) Contribution of intra- and intermolecular hydrogen bonds to the conformational Stability of human lysozyme. *Biochemistry* **38,** 12698–12708